Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Table of contents:

RV217 Study Team and Affiliations

Elizabeth M. Adams[1,2], Siriwat Akapirat[3], Fatim Cham[1,4,5], Agnes Chenine[1,4], Angela Eser[6], Alan Fix[7], Christof Geldmacher[8,9], Michael Hoelscher[8,9], Matthew Johnston[1,3,4], Gustavo H. Kijak[1,4], Shelly J. Krebs[1,4], Arne Kroidl[8,9], Inge Kroidl[6,8,9], Anjali Kunz[10], Cornelia Lueer[6], Leonard Maboko[6], Susan Mason[1,4], Monica Millard[4,5], Joseph Oundo[11], Victoria R. Polonis[4], Erica Sanga[6], Alexandra Schuetz[1,3,4], Arthur Sekiziyivu[12], Doug N. Shaffer[11*], Kibet Shikuku[11], Bonnie H. Slike[1,4], Hendrik Streeck[4,13,], Don Stablein[14], Ali Taylor[1,4], Rapee Trichavaroj[3], Aimee Valenzuela[15], Amy Weintrob[16]

Affiliations
[1]Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD,
[2]Division of AIDS, NIAID, NIH, Bethesda, MD,
[3]Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand,
[4]U.S. Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, MD,
[5]Makere University Walter Reed Project, Kampala, Uganda,
[6]National Institute for Medical Research-Mbeya Medical Research Centre, Mbeya, Tanzania,
[7]PATH, Washington, DC,
[8]German Center for Infection Research (DZIF), partner site Munich, Germany,
[9]Division of Infectious Diseases and Tropical Medicine, Medical Center of the University of Munich (LMU), Munich, Germany,
[10] Madigan Army Medical Center, JBLM, Tacoma, WA
[11]Walter Reed Project, Kericho, Kenya,
[12]Infectious Diseases Institute-Makerere University College of Health Sciences, Kampala Uganda,
[13]Institute for HIV Research, University Hospital Essen, University Duisburg-Essen, Essen, Germany,
[14]EMMES Corporation, Rockville, MD,
[15]Mathematica Policy Research, Washington, DC,
[16] George Washington University, Washington, DC and Veterans Affairs Medical Center, Washington, DC,

* Current Affiliation-Office of the Global AIDS Coordinator, Department of State, Washington, DC

Additional Study Details: RV 217 "HIV-1 Prevalence, Incidence, Cohort Retention, and Host Genetics and Viral Diversity in Cohorts in East Africa and Thailand"

Design Details:

Phase Ia: Non-randomized, cohort, prospective, 24-month observational study to be conducted in two parts. The study conducted a screening visit and initial follow-up in all enrolled volunteers. This included both HIV negative and HIV prevalent cases. Prevalence was estimated to be high in these populations and as many as 1000-1500 volunteers will be enrolled and evaluated in the first two visits with an enrollment target of 500 HIV negative, high-risk volunteers per site. Subsequently, only HIV negative volunteers were followed for a period of 24 months, except for a small number of HIV positive individuals for masking to minimize risk of stigmatization. After the initial two visits, volunteers were seen at the research clinic approximately every 3 months. In addition, at the research clinic or other locations that were convenient to the volunteers, small blood volume samples were collected using microvettes twice weekly. These samples were analyzed and results returned to the site within 48 hours to permit new HIV positive participants to be identified while they remain antibody negative, HIV nucleic acid test positive. Any participant with a reactive nucleic acid test was entered into phase Ib.

Phase Ib: Participants had repeat medical encounters twice weekly for a total of 9 visits to include HIV-1 nucleic acid testing, HIV antibody testing and larger blood collections to characterize host immune responses and a structured history and physical. The purpose of this intensive analysis was to establish HIV diagnosis and characterize the viral-host interaction during acute HIV infection. All newly infected participants were referred to phase II of the study.

Phase II: study of HIV incident cases arising within phase I. HIV incident volunteers were asked to consent to participate in phase II to intensively characterize viral dynamics, host immune response and potential impact of host genetics over a period of at least 50 months from the time of entering the AHI phase..

Participants

The study population will include men and women, aged 18-50 years old, who may be members of the following high risk groups: female and male sex workers (SWs), barworkers (BWs), sexually transmitted infection clinic attendees (STIAs), motorcycle passenger transporters (Boda-Boda), transgenders (TG), and men who have sex with men (MSM). Although these populations are known to be among "Most at-Risk Populations" or MARPs, not all members of these occupational groups engage in equivalent risk behavior. Enrollment will therefore engage members of these populations whose behavior places them at higher risk as documented through an Audio-Computer Assisted Self-Interview. Only participants identifying at least one of four risk factors were enrolled (see table S1).

Supplementary methods:

Informed Consent: Written informed consent was obtained in the native language of the participant. Standard procedures for enrolling illiterate volunteers were observed.

HIV serological diagnostics:
HIV serology was performed at screening, every 6 months and during phase IB using the Genetic Systems HIV-1/HIV-2 Plus O EIA (Bio-Rad Laboratories, Inc; Redmond, WA). Reactive samples were repeated in duplicate then confirmed by the Genetic Systems HIV-1 Western Blot (Bio-Rad Laboratories).

HIV subtyping: HIV-1 near full-length genomes or two overlapping half genomes were sequenced from plasma RNA using a single genome amplification strategy.[1] The PCR amplification and sequencing methods were as described.[2] HIV-1 subtype was assigned using the HIV-1 Genotyping Tool at the National Center for Biotechnology Information and confirmed by Maximum Likelihood phylogenetic analysis and *jumping profile hidden Markov model* (*jpHMM*).[3,4]

Spline statistical methods:
The spine was defined using the truncated power function basis with degree 2. The number of spline knots were selected by minimizing the Akaike information criteria (AIC). Region-specific spline effects were were added to the base model after being tested for significance using the AIC. Subject-specific splines were also considered but did not improve model fit or the AIC. Splines models were fit in PROC GLIMMIX.

Flow Cytometry Analysis:
Lymphocyte analysis focused on visits where more than 30 observations were available. Models were fit using restricted maximum likelihood estimation accounting for within subject covariance using unstructured covariance matrices for univariable and multivariable models. All models were adjusted for region and visit.

Study Conduct and Analysis: The study was designed by Merlin Robb, Leigh Anne Eller, Sheila Peel, Robert O'Connell, Mark deSouza, and Nelson Michael. Data were gathered by Hannah Kibuuka, Kathleen Rono, Lucas Maganga, Sorachai Nitayaphan, Eugene Kroon, Somchai Sriplienchan, Samuel Sinei, Fred Sawe, Linda Jagodzinski, Mark Manak, Sodsai Tovanabutra, and Mike Eller. Data analysis was performed by Merlin Robb, Leigh Anne Eller, Andrew Lewandowski, Hao Wu, and Peter Dawson. Merlin Robb, Leigh Anne Eller, Mark Millazo, and Peter Dawson vouch for the data. The paper was written by Merlin Robb, Leigh Anne Eller, Peter Dawson, Jerome Kim and Nelson Michael. The decision to publish the paper was made by Merlin Robb and Nelson Michael.

Table S1: Description of screening, surveillance and acute infection cohorts

	All Sites	Kenya	Tanzania	Uganda overall	Uganda-Female	Uganda MSM	Thailand Overall	Thailand FSW	Thailand MSM	Thailand TG
SCREENING POPULATION										
Volunteers screened	3954	1046	1010	870	681	187	1028	165	444	418
HIV prevalence at screening (%)	28.4	23.0	40.3	32.2	37.0	14.2	19.5	1.3	28.2	17.7
Entry risk criteria for high risk pop										
Received goods for sex (%)	64	80	18	68	70	61	87	88	87	88
Unprot. Sex w/ known HIV+ (%)	17	4	29	14	11	22	19	9	24	18
Unprot Sec w/ 3 or more (%)	50	50	34	66	61	88	51	27	58	53
Reported STI symptoms (%)	61	36	86	69	71	58	54	76	51	49
Lost to follow up once in SBV (%)	30.0	28.4	34.0	29.5	35.1	14.1	29.2	30.5	28.6	30.3
SBV compliance (%)	86.8	95.4	84.7	89.7	90.2	87.9	77.2	81.8	77.6	74.1
ACUTE INFECTIONS										
Total # acute cases	112	28	21	13	9	4	50	2	32	16
Incidence rate (95% CI)	3.3 (2.7-3.95)	2.6 (1.7-3.6)	3.1 (1.8-4.4)	2.0 (0.9-3.1)	1.7 (0.6-2.8)	3.6 (0.7-7.1)	5.1 (3.7-6.5)	0.8 (0.3-2.0)	8.4 (5.5-11.3)	4.5 (2.3-6.7)
Person-Years Follow Up	3362	1060	684	640.3	529	111	977.7	239.8	380.2	357.7
Median age incident cases (range)	23 (18-48)	24 (18-34)	23 (18-34)	24 (18-33)	26 (18-33)	21 (18-24)	23 (18-48)	40 (32-48)	23 (18-35)	22 (18-29)
Entry high risk criteria										
Received goods for sex (%)	77	89	24	85	78	100	90	100	94	81
Unprot. Sex w/ known HIV+ (%)	21	4	29	8	11	0	30	50	31	25
Unprot Sec w/ 3 or more (%)	54	61	43	85	78	100	46	50	47	44
Reported STI symptoms (%)	58	46	91	46	67	0	54	50	59	44

[1]Participants were enrolled if they reported one of the following: 1) vaginal or anal intercourse in exchange for money, goods or services 2) unprotected vaginal or anal intercourse with one or more known HIV-positive partners 3) unprotected vaginal or anal intercourse with three or more partners of known or unknown HIV status 4) new sexually transmitted disease.
STI=Sexually Transmitted Infection; SBV=Small blood volume (twice weekly collections)

Table S2: Acute cohort: Analysis and excluded cohorts

	All cases (112)	n=50 included in primary analysis	n=62 excluded from primary analysis	p value
# Gender (%)	Male: 52 (46) Female: 60 (54)	Male: 19(38) Female: 31(62)	Male: 33(53.2) Female: 29(46.8)	0.13
Median Age (IQR)	23(20-27)	24(20-25)	23(20-27)	0.72
# per site (%)	Kenya: 28 (25) Uganda: 13 (11.6) Tanzania: 21 (18.8) Thailand: 50 (44.6)	Kenya: 16(32) Uganda: 9(18) Tanzania: 8(16) Thailand: 17(34)	Kenya: 12(19.4) Uganda: 4(6.5) Tanzania: 13(21) Thailand: 33(53.2)	0.057
Entry high risk criteria (%)	Received goods for sex: 77 Unprot. Sex w/ known HIV+: 21 Unprot Sec w/ 3 or more: 54 Reported STI symptoms: 58	Received goods for sex: 80 Unprot. Sex w/ known HIV+: 12 Unprot Sec w/ 3 or more: 58 Reported STI symptoms: 50	Received goods for sex: 74 Unprot. Sex w/ known HIV+: 27 Unprot Sec w/ 3 or more: 50 Reported STI symptoms: 65	0.51 0.060 0.45 0.13
Median Viral Load set point ($\log_{10}$ copies/ml)	4.31 (3.80-4.87), n=78	4.35 (3.70-4.83), n=45	4.31 (4.09-4.92), n=33	0.73

Table S3: Summary quantitative viral RNA statistics

	Mean, Median (IQR)		
	All	**Africa**	**Thailand**
Peak VL (n=50)	6.75, 6.74 (6.24, 7.50)	6.76, 6.76 (6.12, 7.53)	6.73, 6.66 (6.26, 6.91)
Peak Day (n=50)	12.5, 13 (10, 14)	12.2, 13 (10, 14)	13.1, 14 (11, 15)
Nadir VL(n=47)[1]	4.24, 4.33 (3.55, 4.79)	4.03, 4.22 (3.43, 4.71)	4.65, 4.70 (4.14, 5.44)
Nadir Day(n=47)[1]	32.7, 31 (28, 41)	32.1, 31 (27, 38)	33.9, 33 (28, 42)
Set-Point VL (n=45)[2]	4.26, 4.35 (3.7, 4.83)	4.00, 3.99 (3.39,4.50)	4.69, 4.77 (4.41, 5.22)
EIA Reactive Day (n=43)[3]	15.4, 14 (13,18)	15.1, 14 (13-18)	15.4, 14 (13,18)
Up Slope (n=43)[4]	0.40, 0.39 (0.35, 0.43)	0.41, 0.40 (0.36, 0.44)	0.38, 0.38 (0.34, 0.41)
Down Slope (n=49)[5]	-0.14, -0.13 (-0.16, -0.09)	-0.15, -0.14 (-0.18, -0.10)	-0.11, -0.10 (-0.13, -0.08)

[1] 2 cases not included due to start of ART and 1 case not included due to missing visits
[2] 5 cases not included due to start of ART
[3] 7 cases not included due to missing data
[4] 7 cases not included due to interval >10 days between last negative and first positive
[5] 1 case not included due to missing visit

Figure S1: Viral Load Regional Spline Models

A.
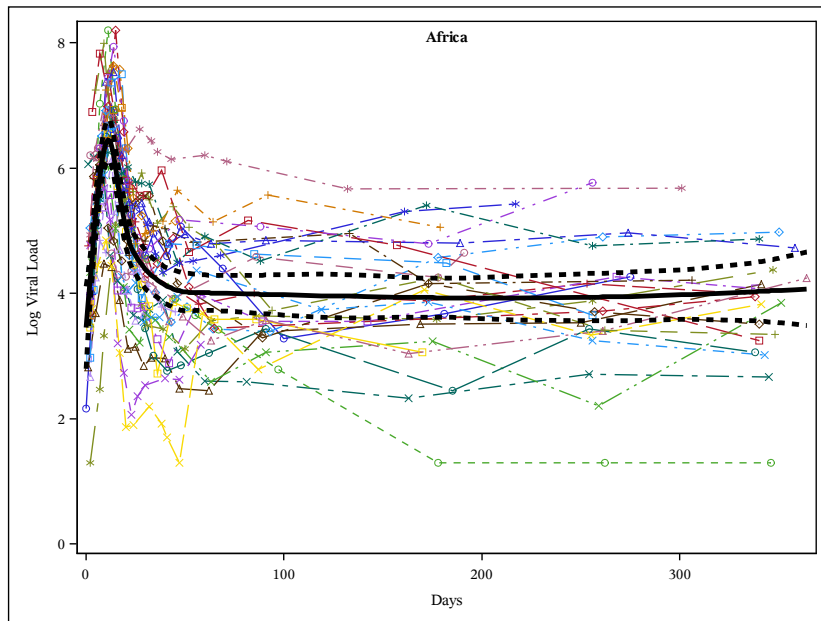
B.
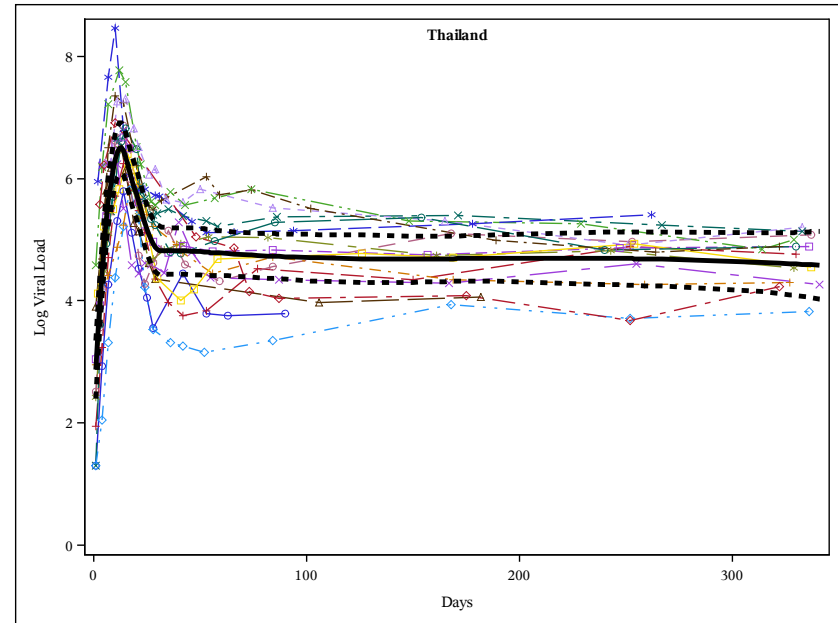
Figure S1: Regression splines with subject-specific slopes and intercepts were fit using the first year of viral load data for each region (panel B and C). Day 0 is defined as the first reactive APTIMA RNA test. The spline was defined using the truncated power function basis with degree 2. The number of spline knots were selected by minimizing the Akaike information criteria (AIC). Region-specific spline effects were added to the base model after being tested for significance using the AIC. Subject-specific splines were also considered, but did not improve model fit or the AIC. Splines models were fit in PROC GLIMMIX of SAS version 9.3.

Table S4: Individual Viral Load Data

Region	peak day	peak vl log$_{10}$ copies/ml	nadir day log$_{10}$ copies/ml	nadir vl log$_{10}$ copies/ml	set point vl log$_{10}$ copies/ml	Day of last negative	VL Downslope (Rate of Change Peak to Nadir)	VL Upslope (Rate of Change Last Negative to Peak)	Last known Absolute CD4 count (cells/ul)	Absolute CD4 Count 1 year (cells/ul)
Africa	11	7.51	34	4.29	3.97	-21	-0.14		556	771
Africa	12	5.57	42	2.88	3.89	-28	-0.09		594	594
Africa	18	5.49	39	3.43	3.69	-4	-0.10	0.25	618	1035
Africa	10	6.92	35	4.01	4.08	-5	-0.12	0.46	290	514
Africa	14	6.74	42	3.55	3.52	-14	-0.11		746	746
Africa	13	6.62	27	4.59	4.13	-14	-0.15		625	554
Africa	6	6.32	30	4.22	4.35	-162	-0.09		453	
Africa	14	7.63	31	4.94		-4	-0.16	0.42	442	
Africa	14	7.94	32	4.71	5.20	-4	-0.18	0.44	323	323
Africa	9	7.99	36	5.12	4.28	-4	-0.11	0.61	1543	1412
Africa	8	6.25	28	3.57	2.66	-32	-0.13		376	579
Africa	11	8.2	31	4.55	2.52	-4	-0.18	0.55	814	873
Africa	13	6.68	29	3.54	3.70	-4	-0.20	0.39	345	347
Africa	14	7.64	25	4.79	5.35	-4	-0.26	0.42	495	265
Africa	11	6.02	27	3.21	3.01	-5	-0.18	0.38	368	760
Africa	13	6.12	25	3.56		-4	-0.21	0.36	676	
Africa	11	6.54	36	2.72	3.57	-4	-0.15	0.44	788	
Africa	13	6.95	39	4.33	4.87	-4	-0.10	0.41	186	186
Africa	15	8.2	31	4.31	4.02	-4	-0.24	0.43	725	890
Africa	14	6.93	41	2.76	3.04	-4	-0.15	0.39	534	661
Africa	14	6.73	22	4.61	4.54	-4	-0.27	0.37	249	257
Africa	9	5.68	23	2.06	2.90	-21	-0.26		465	
Africa	9	4.47	29	2.84	3.25	-4	-0.08	0.34	669	1041
Africa	17	5.45	38	3.78		-7	-0.08	0.23	935	935
Africa	12	7.57				-3		0.50	509	
Africa	18	7.5	28	3.93	4.38	-3	-0.36	0.36	427	618
Africa	14	6.76	32	5.75	4.83	-3	-0.06	0.40	382	454
Africa	11	7.34	36	6.26	5.96	-10	-0.04	0.35	474	474
Africa	10	4.85	41	1.7	3.16	-3	-0.10	0.37	393	685
Africa	7	6.05	18	4.52	3.77	-10	-0.14	0.36	259	463
Africa	14	7.53	42	4.81	4.85	-4	-0.10	0.42	379	379
Africa	7	7.83	27	5.56	4.46	-7	-0.11	0.56	1018	1021
Africa	18	6.96				-3		0.33	583	
Thailand	15	7.31	42	5.61	5.35	-3	-0.06	0.41	360	421
Thailand	14	5.79	28	3.55	3.78	-3	-0.16	0.34	307	
Thailand	16	6.4	41	4	4.64	-7	-0.10	0.28	180	627
Thailand	10	8.46	31	5.71	5.22	-7	-0.13	0.50	180	180
Thailand	14	6.24	42	3.76	4.47	-5	-0.09	0.33	249	868
Thailand	11	6.66	42	5.38	5.28	-6	-0.04	0.39	283	464
Thailand	13	6.5	29	4.36	4.02	-4	-0.13	0.38	695	
Thailand	14	6.67	35	4.79	4.83	-5	-0.09	0.35	491	462
Thailand	10	7.35	28	5.49	5.39	-3	-0.10	0.57	189	264
Thailand	12	6.49	26	4.62	4.83	-2	-0.13	0.46	343	358
Thailand	17	5.81	42	3.26	3.60	-4	-0.10	0.28	490	433
Thailand	14	6.86	24	4.28	4.77	-5	-0.26	0.36	225	391
Thailand	18	6.09	42	4.78	4.41	-2	-0.05	0.30	257	603
Thailand	10	6.91			4.30	-7		0.41	781	971
Thailand	7	6.26	21	4.45	4.45	-9	-0.13	0.39	431	525
Thailand	15	6.82	41	4.77	5.07	-4	-0.08	0.36	154	356
Thailand	12	7.77	29	5.57	5.35	-7	-0.13	0.41	81	344

Figure S2: Regional viral load correlations
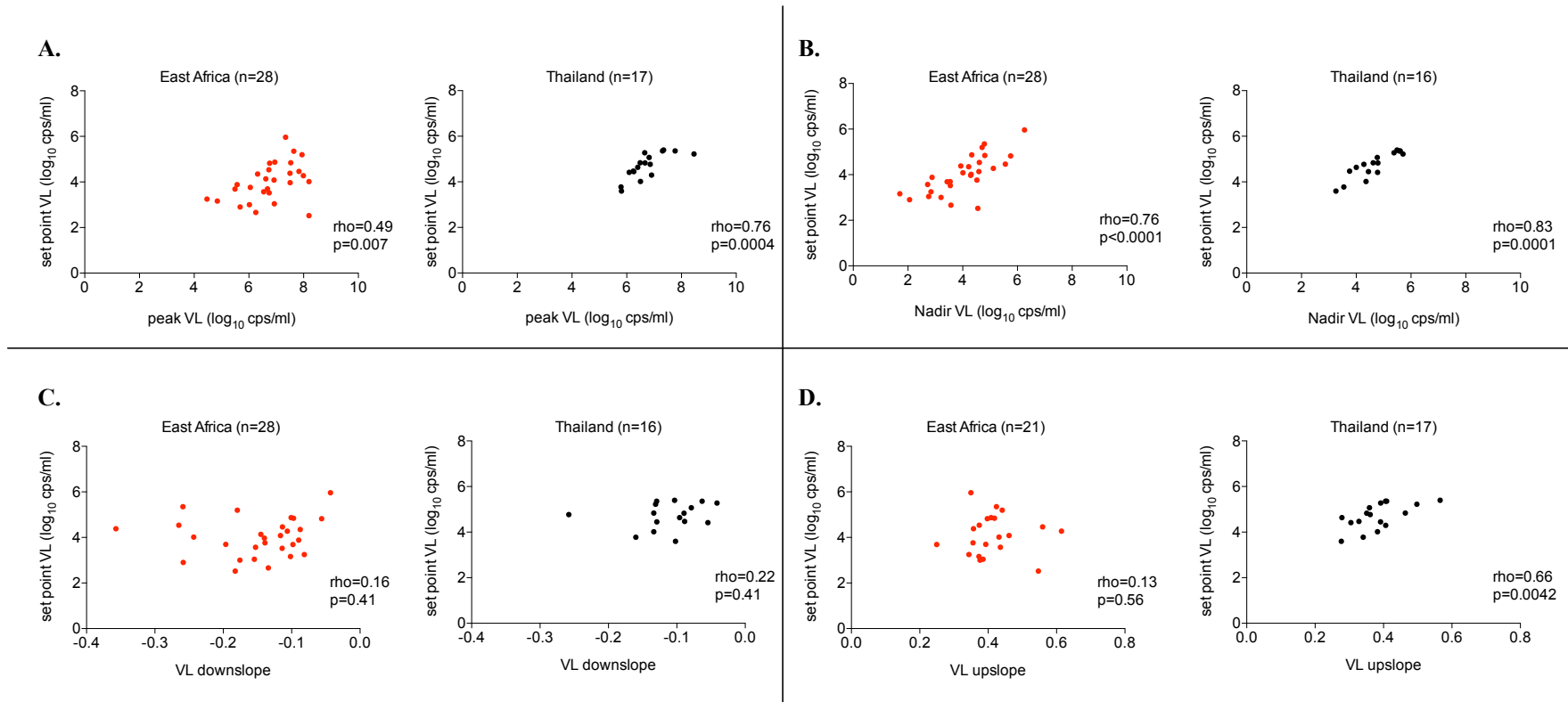
Figure S2: Peak vial load correlated to set point viral load in East Africa (EA) and Thailand (TH) (Panel A). Nadir and set point viral load were strongly correlated in both regions (Panel B). Although downslope correlated to set point in the combined group (Figure 3D, main paper), when separated by region, the correlation was not present in East Africa (EA) or Thailand (Panel C). The correlation of viral load upslope to set point viral load was only significant in Thailand (Panel D).

Figure S3: Immune cell counts over 510 days of follow up by region
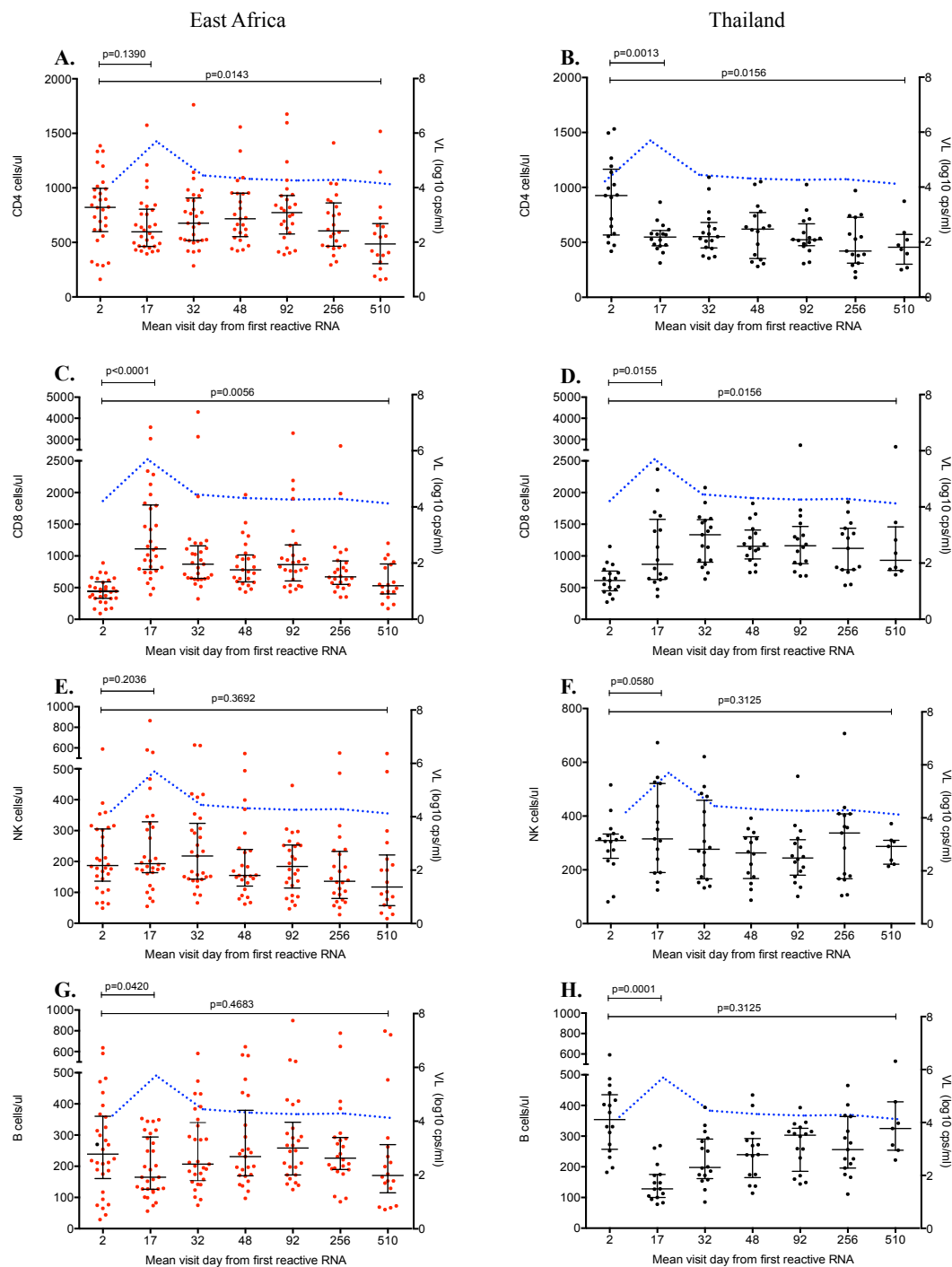
Figure S3: The absolute counts for CD4 (Panel A,B), CD8 (Panel C,D), NK cell (Panel E,F) and B cell (Panel G,H) for 50 subjects were plotted by the mean visit day on the x-axis. African participants (n=32) are shown on the left and Thailand participants (n=18) on the right. The mean viral load at each time point is plotted in blue on the right Y-axis. Wilcoxon Signed-Rank tests were performed on available paired data prior to start of ARV treatment.

Table S5: Frequency of clinical symptoms and signs overall and by region

	Africa (n=31)	Thailand (n=17)	All (n=48)
Symptom:	N(%)	N(%)	N(%)
Fever	18 (55%)	7 (41%)	25 (50%)
Headache	17 (52%)	6 (35%)	23 (46%)
Feeling of Illness	14 (42%)	5 (29%)	19 (38%)
Coughing	10 (30%)	9 (53%)	19 (38%)
Abnormality:			
HEENT**	6 (18%)	16 (94%)	22 (44%)
Lymph*	3 (9%)	16 (94%)	19 (38%)
Tachycardia	11 (33%)	5 (29%)	16 (32%)

*Fisher's Exact p-value < 0.001
**Fisher's Exact p-value <0.001

Figure S4A: Heat map-Symptoms reported in medical history

Region	Pre-Infection	Visit 5 Pre-Peak	Visit 4 Pre-Peak	Visit 3 Pre-Peak	Visit 2 Pre-Peak	Visit 1 Pre-Peak	Peak	Visit 1 After Peak	Visit 2 After Peak	Visit 3 After Peak	Visit 4 After Peak	Visit 5 After Peak	Total # Visits Post-infection w/ Symptoms	Total Reported Symptoms Post-infection
Africa	0	.	.	0	0	0	0	0	0	0	0	0	0	0
Africa	0	.	0	0	0	0	0	0	.	0	0	0	0	0
Africa	0	.	.	.	0	0	0	0	0	0	0	0	0	0
Africa	0	.	.	.	0	0	0	0	0	0	0	0	0	0
Africa	0	.	.	0	0	0	0	0	0	0	0	0	0	0
Thailand	0	.	0	0	0	0	0	0	0	0	0	0	0	0
Africa	0	.	.	.	6	.	0	0	0	0	0	0	1	6
Africa	0	.	.	.	.	3	0	0	0	0	0	0	1	3
Africa	0	.	.	0	0	.	0	0	0	0	0	4	1	4
Africa	0	.	.	0	3	0	0	0	0	0	0	0	1	3
Africa	0	.	.	0	3	0	0	0	0	0	0	0	1	3
Africa	0	.	.	0	1	0	0	.	0	0	0	0	1	1
Africa	0	.	.	.	0	1	0	0	0	0	0	0	1	1
Africa	0	.	0	4	0	0	0	0	0	0	0	0	1	4
Africa	0	.	.	.	.	0	0	0	0	4	0	0	1	4
Thailand	0	.	.	.	.	2	0	0	0	0	0	0	1	2
Africa	0	.	.	0	6	1	0	0	.	0	0	0	2	7
Africa	0	.	.	.	1	0	2	0	0	0	0	.	2	3
Africa	1	.	.	.	1	1	0	0	0	0	0	0	2	2
Africa	3	.	.	.	9	0	0	0	0	0	0	1	2	10
Africa	0	.	.	.	0	9	1	0	0	0	0	0	2	10
Africa	0	.	.	0	0	11	3	0	0	0	0	0	2	14
Africa	0	1	0	1	0	0	0	0	0	0	.	0	2	2
Africa	0	.	.	.	.	5	8	.	.	0	0	0	2	13
Africa	0	.	4	.	2	0	0	.	.	.	.	.	2	6
Thailand	0	.	0	0	0	1	1	0	0	0	0	0	2	2
Thailand	2	.	.	.	0	1	2	0	0	0	0	0	2	3
Thailand	0	.	.	0	0	0	3	1	0	0	0	0	2	4
Thailand	6	.	.	.	0	7	3	0	0	.	0	0	2	10
Thailand	0	.	.	0	0	0	1	1	0	0	0	0	2	2
Africa	0	.	.	.	0	2	17	0	1	0	0	0	3	20
Africa	0	.	.	0	2	1	5	.	.	.	.	.	3	8
Africa	0	.	.	0	1	4	0	0	0	0	0	7	3	12
Thailand	0	.	.	.	1	3	5	.	.	.	.	.	3	9
Africa	0	.	.	.	5	0	0	0	3	0	1	0	3	9
Africa	0	.	.	2	0	1	2	2	0	0	0	0	4	7
Africa	3	.	.	0	0	15	2	0	2	0	1	0	4	20
Africa	1	.	.	.	2	1	4	0	1	0	0	0	4	8
Thailand	1	.	2	0	0	4	1	0	0	1	0	0	4	8
Thailand	0	3	2	1	0	0	0	2	0	0	0	0	4	8
Thailand	0	.	0	0	0	6	3	2	0	2	0	0	4	13
Thailand	0	.	.	0	4	2	1	1	1	0	0	0	5	9
Thailand	0	1	0	0	2	3	1	0	1	0	0	0	5	8
Africa	0	.	.	.	.	5	10	1	0	2	0	1	5	19
Africa	0	.	.	.	5	2	3	1	1	0	2	0	6	14
Africa	0	.	.	.	0	1	2	1	1	0	1	1	6	7
Thailand	0	.	1	1	0	1	1	1	1	0	1	0	7	7
Thailand	0	.	1	0	4	2	1	0	1	1	1	0	7	11
Thailand	0	.	2	2	3	1	0	0	1	2	1	1	8	13
Thailand	0	.	.	0	1	7	1	4	1	1	1	1	8	17
Average # Symptoms per Visit	0.34	1.7	0.92	0.39	1.4	2.1	1.7	0.38	0.34	0.28	0.2	0.35		

13

Figure S4B: Heat map-Physical Exam findings

Region	Pre-Infection	Visit 5 Pre-Peak	Visit 4 Pre-Peak	Visit 3 Pre-Peak	Visit 2 Pre-Peak	Visit 1 Pre-Peak	Peak	Visit 1 After Peak	Visit 2 After Peak	Visit 3 After Peak	Visit 4 After Peak	Visit 5 After Peak	Total # Visits Post-infection w/ Physical Abnormalities	Total Physical Abnormalities Post-Infection
Africa	0	.	.	0	0	0	0	0	0	0	0	0	0	0
Africa	0	.	0	0	0	0	0	0	.	0	0	0	0	0
Africa	0	.	.	.	0	0	0	0	0	0	0	.	0	0
Africa	0	.	.	.	0	0	0	0	0	0	0	0	0	0
Africa	0	.	.	.	.	0	0	0	0	0	0	0	0	0
Africa	0	.	.	0	0	.	0	0	0	0	0	0	0	0
Africa	0	.	.	.	0	0	0	0	0	0	0	0	0	0
Africa	0	.	.	.	.	0	0	0	0	0	0	0	0	0
Africa	0	.	.	0	0	0	0	0	0	0	0	0	0	0
Africa	0	.	.	.	.	0	0	.	.	0	0	0	0	0
Africa	0	.	0	.	0	0	0	.	.	.	.	.	0	0
Africa	0	.	.	0	1	0	0	0	.	0	0	0	1	1
Africa	0	.	.	.	1	.	0	0	0	0	0	0	1	1
Africa	0	.	.	.	0	0	0	1	0	0	0	0	1	1
Africa	0	.	.	.	0	0	0	0	1	0	0	0	1	1
Africa	1	.	0	3	0	0	0	0	0	0	0	0	1	3
Africa	0	.	.	0	0	1	0	0	0	0	0	0	1	1
Thailand	0	.	0	0	1	0	0	0	0	0	0	0	1	1
Africa	0	.	.	0	0	0	1	1	0	0	0	0	2	2
Africa	0	.	.	1	1	0	0	0	0	0	0	0	2	2
Africa	0	.	.	1	0	2	0	0	0	0	0	0	2	3
Africa	0	.	.	0	0	1	1	.	.	.	.	.	2	2
Africa	0	0	0	1	0	0	0	0	0	0	.	1	2	2
Africa	0	.	.	.	0	0	0	0	1	0	1	0	2	2
Africa	1	.	.	.	.	1	2	0	0	1	0	0	3	4
Africa	0	.	.	.	1	1	0	1	0	0	0	0	3	3
Africa	1	.	.	.	0	2	1	1	0	0	0	0	3	4
Thailand	1	.	.	.	2	3	3	.	.	.	.	.	3	8
Africa	0	.	.	0	1	0	0	0	1	1	0	0	3	3
Africa	0	.	.	.	2	0	2	1	1	0	0	0	4	6
Thailand	0	.	1	0	0	1	1	0	1	0	0	0	4	4
Thailand	0	.	1	0	0	2	2	0	0	0	0	1	4	6
Thailand	0	1	0	0	2	1	0	0	0	1	0	0	4	5
Africa	0	.	.	0	0	3	1	2	1	0	1	0	5	8
Thailand	0	.	1	0	0	1	2	2	1	0	0	1	6	8
Thailand	1	.	.	.	1	1	2	1	1	.	1	0	6	7
Africa	1	.	.	.	3	2	0	0	3	2	1	1	6	12
Africa	0	.	.	.	2	1	2	1	3	1	0	0	6	10
Africa	0	.	.	.	2	0	2	1	1	1	2		7	10
Thailand	0	.	.	.	2	1	2	1	1	1	0	1	7	9
Thailand	1	.	.	.	.	1	1	1	1	1	1	1	7	7
Africa	0	.	.	1	1	2	1	2	2	0	2	2	8	13
Thailand	0	.	0	0	1	2	2	1	2	2	2	1	8	13
Thailand	0	.	0	2	0	2	2	2	1	1	2	1	8	13
Thailand	1	.	.	1	2	1	2	2	2	1	1	1	9	13
Thailand	0	.	.	1	1	1	1	1	1	1	1	1	9	9
Thailand	1	.	.	1	2	2	2	2	1	1	1	2	9	14
Thailand	0	.	.	1	1	1	1	1	1	2	1	1	9	10
Thailand	2	.	1	3	2	2	1	1	1	2	2	2	10	17
Thailand	0	2	2	2	1	1	1	2	2	1	1	1	11	16
Average # of Abnormalities per visit	0.22	1	0.46	0.64	0.73	0.81	0.76	0.61	0.68	0.43	0.41	0.43		

Figure S4A and S4B: The number of symptoms reported in the medical history (Figure S4A) and physical findings on exam (Figure S4B) at each visit is shown for each participant. The visit for each participant where peak viral load occurred is shown at the center of the table and the visits preceding peak viral load are to the left and visits subsequent to peak viral load are on the right. The Pre column indicates the number of reported symptoms for each participant at a pre-infection visit. The total number of findings and visits with findings are shown on the right for each participant.

References:

1.	Salazar-Gonzalez JF, Bailes E, Pham KT, et al. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. Journal of virology 2008;82:3952-70.
2.	Heipertz RA, Jr., Sanders-Buell E, Kijak G, et al. Molecular epidemiology of early and acute HIV type 1 infections in the United States Navy and Marine Corps, 2005-2010. AIDS research and human retroviruses 2013;29:1310-20.
3.	Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T. A web-based genotyping resource for viral sequences. Nucleic acids research 2004;32:W654-9.
4.	Schultz AK, Zhang M, Bulla I, et al. jpHMM: improving the reliability of recombination prediction in HIV-1. Nucleic acids research 2009;37:W647-51.